Trusting versus Monitoring: An Experiment of Endogenous Institutional Choices °

Andrej Angelovski[*], Daniela Di Cagno[*], Daniela Grieco[♣], Werner Güth[*♦]

29.03.2019

Abstract

We investigate the problem of deciding between trusting and monitoring, and how this decision affects subsequent behavior, using a laboratory experiment where subjects choose between the Ultimatum and the Yes-No Game. Despite the similarity of the two games in Ultimatum Games responders monitor the allocation proposal, while in Yes-No games responders react without monitoring, i.e. have to rely on trust. We permit either the proposer or responder to make the game choice and analyze how both roles choose between trusting and monitoring, what the ensuing effects of their choices are, and how they vary depending on who has chosen the game. We, also, experimentally vary the cost of monitoring and the responder's conflict payoff. Since monitoring is usually costly, the amount to share in Yes-No Games (YNG) can exceed that in Ultimatum Games (UG). Regarding the conflict payoff, it can be positive or negative with the former rendering Yes-No interaction a social dilemma. According to our results, proposers (responders) opt for trusting significantly more (less) often than for monitoring. Average offers are higher in Ultimatum than in Yes-No games, but neither UG nor YNG offers depend on who has chosen between games.

Keywords: Monitoring, Trusting, Institutional Choice, Ultimatum Game
JEL: C91, C92, M52

° The authors gratefully acknowledge the constructive advice of their two anonymous referees.
[*] LUISS Guido Carli, Department of Economics and Finance, Rome.
[♣] Bocconi University, Department of Economics, Milan.
[♦] Max Planck Institute for Collective Goods, Bonn.

This research project was financed by the Max Planck Institute for Research on Collective Goods.
Corresponding Author: Andrej Angelovski (aangelovski@luiss.it)

1. Introduction

In organization management, the trade-off between monitoring and trusting arises quite often. It refers to the possibility to better incentivize employees to behave properly. In addition, monitoring performance and compliance of partners in joint venture firms is a crucial aspect of corporate governance. Monitoring often implies suspicion and mistrust and is usually rather costly, both materially and emotionally. If trusting each other would be the norm, embarrassing practices like checking employee email and internet use, examining business expenditures, inspecting private belongings upon exit, unannounced performance evaluations etc. could be avoided, which would lower overall costs and make our professional life more enjoyable.[1] Individuals can, however, behave "badly" and thus anticipated monitoring and sanctioning may have to be used as a prevention or sanctioning tool. Being aware of the possibility of "bad" behavior, even trustees themselves, for instance employees, might prefer monitoring.

Whether joint venture partners engage in mutual monitoring and whether an employer monitors employee performance is, therefore, an important aspect of human resource management (see, for instance, de Kok and Uhlaner, 2001; Audretsch and Thurik, 2000 and 2001). When "an atmosphere of mutuality and trust" indeed generates higher productivity gains and lower unit costs (claimed, for instance, by de Kok and Uhlaer, 2001; and Boxall, 1996), trusting rather than monitoring may be the better way of managing human resources as it enhances intrinsic work motivation and identification with the firm. Monitoring can crowd out intrinsic work motivation, question loyalty, and lead to corporate identity concerns.[2] Trust may therefore act as a "substitute for formal contracts" and avoid completing corporate governance by auditing routines and contract clauses related to monitoring (Hendrikse et al., 2015, discuss how trusting relates to incompleteness of contracts). In light of this, and considering the potentially high monetary costs related to monitoring, such as acquiring and maintaining monitoring systems or hiring additional supervisors, the debate on when to trust and when to monitor will likely continue.

This paper investigates the determinants of individual (or corporate) decisions on whether to trust or to monitor and not on how to monitor, what in large commercial enterprises is often highly regulated. Nevertheless, the setting in which the decision between trusting and monitoring is made may also play a crucial role. Monitoring is likely preferred when firms are large, social interaction is more anonymous and stakes are larger, while small firms with few employees, who repeatedly interact in

[1] Frey (1993) stresses the rivalry between trust and loyalty in shaping work effort.

[2] Crowding-out of intrinsic motivation has been studied experimentally (e.g. Falk et al., 1999; Nagin et al., 2002; Falk and Kosfeld, 2006) and theoretically (e.g. Benabou and Tirole, 2003; Bowles and Polania-Reyes, 2012).

person, may be more likely to trust each other while.[3] Trusting in small firms, however, may still depend on whether individual or team efforts are hidden and on their private information (see, for instance, Halac and Pratt, 2016, who discuss ways how to implement monitoring). In such firms, one may continuously and directly check the quality of their employees' performance or, alternatively, trust them.

Given that a general and conclusive answer on what is the best choice is quite hard to answer, one may still want to discern some determinants of the decision experimentally. Some experimental research has explored how monitoring and trust interact (Dickinson and Villeval, 2008; Schweitzer et al., 2016) although, for the most part, the relationship still remains to be disentangled. Using a real effort task, Dickinson and Villeval vary monitoring intensity in a principal/agent setting and show that more monitoring increases performance (see also Calvo and Wellisz, 1978; Fama and Jensen, 1983; Prendergast, 1999). They find that principals prefer to monitor more, which enhances the performance of most trustees (agents), although above a certain level of monitoring crowding-out occurs. The dynamics of monitoring or trusting are analyzed by Schweitzer et al. (2016) via a repeated trust game with varying monitoring conditions: trustees are more trustworthy when anticipating monitoring and less trustworthy when not, however, trustors often fail to anticipate this conditioning and continue to trust without monitoring.

Exogenously imposed trusting or monitoring avoids an endogenous selection with whom one interacts: trustees, who opt for trusting, may cause suspicion[4] whereas trustors opting for it may want to signal trust in the trustworthiness of their trustees. While there are many examples in private and economic environments with one of the interacting parties deciding between monitoring or trusting others' performance, only a few studies focus on the endogenous choice between trusting and monitoring by either of the involved parties. Although the effect of trust on organizations and group performance has been widely studied[5], the determinants which prompt one to choose trusting are less explored. In addition to the endogenous selection by one party, the effect of who, a trustor or trustee,

[3] An example of the latter case is the International Atomic Energy Agency (IAEA), in Vienna, which monitors the compliance of many member countries of the Non-Proliferation Treaty for Nuclear Weapons (NTP). Since nuclear energy provision produces uranium in random amounts, underreporting of uranium is discouraged by randomly inspecting nuclear power plants. By monitoring one wants to limit the risk of setting aside uranium for nuclear weapons. How to randomly inspect and prevent underreporting of uranium amounts is far from obvious (see Avenhaus et al. 1996).

[4] As not monitoring allows for "moral wiggle room" which often is exploited (see Dana et al., 2007).

[5] Some determinants which have been looked at are: individual autonomy (Langfred, 2004), view of one's vs. other's trustworthiness (Ferrin et al., 2007), ongoing (long-term) teams (De Jong and Elfring, 2010), trust in team members vs. trust in supervisors (Bijlsma-Frankema et al., 2008), the role of trust in organizational setting (Dirk and Ferrin, 2001 and Gätcher and Falk (2002).

decides between the institutional options is still unresolved. [6] Thus, we have designed a lab experiment which allows us to focus on the choice between trusting and monitoring by proposer and responder participants without any prior experience: we ask who chooses which institution, and how subsequent behavior is affected by who has made the institutional choice?

We hope to capture the core aspect of the conflict between trusting and monitoring, not by inspection games (see Avenhaus, 2004), but by employing basic paradigms which may be applicable to all such decisions may they occur in business, politics, or private interaction. Specifically, we employ the Ultimatum and the Yes-No Game. Whereas in Ultimatum Games (henceforth, UG) responders monitor the allocation proposal before deciding to accept or reject, in Yes-No games (henceforth, YNG) responders react without monitoring offers while still maintaining veto power. Both games are easily understood, limit demand effects and allow to compare data for either institution with experiments without preceding game selection. Most importantly, we believe that deciding between the two games truly captures the essence of "trusting vs. monitoring".

Trustors, in our setup, are responders and might represent employers choosing between trusting or monitoring their employees[7]. There may also be instances when trustees make this decision, for example, to avoid doubts concerning their performance. We question whether game playing behavior depends on who has selected the game and whether the superior efficiency of YNG (which can create a social dilemma) affects game choice and its game play. We vary these parameters across successive rounds during which the same partners interact repeatedly without receiving feedback on previous outcomes till the end.[8]

UG and YNG both allow for altruistic punishment. In particular, YNG captures trust like in Berg et al. (1995) whereas UG offers are monitored before accepting or rejecting them. Monitoring costs can increase the amount to be shared in YNG compared to UG.

In employment relationship it is usually the employer who is responsible for the cost of monitoring, however, the incidence (burden), i.e. who suffers more from the cost of monitoring, is less obvious (for the related literature on the incidence of tax payments, see Fullerton and Metcalf, 2002). If monitoring costs also capture the psychological costs of being monitored, which depending on the

[6] See Malhotra (2004) for a general discussion of differing perspectives of trustors and trusted parties in traditional trust games.

[7] In corporate governance employers are usually responsible for institutional rules, and are represented by responders in our setup. Proposers, in our setup, represent employees, who may shirk, concisely underperform, steal etc. and thereby lower the amount to be shared.

[8] In repeated interaction with more feedback it can make sense to study whether and how trust can be repaired after its violation, for example, by apologizing or attributing the violation to some external cause (see, for instance, Kim et al., 2006).

situation may range from low to high, employees might be the ones who suffer. The advantage of using two bargaining games is that one game, YNG, avoids monitoring costs and thus increases what both parties can share.

Game-theoretically, both games predict exploitation for non-positive conflict payoffs of responders but, unlike YNG, UG has multiple equilibria in weakly dominated strategies. Only when the responder's conflict payoff exceeds the lowest possible offer, YNG becomes a social dilemma game for common(ly known) opportunism (the responder rejects the unknown offer which would be lowest when the proposer expects acceptance).

Truly intrinsically generous proposer participants should not mind monitoring since anyhow offering a fair share (see Cooper and Kagel, 2016). This may render responders suspicious when confronting a proposer who has opted for YNG.

UG is experimentally implemented by letting responders choose an acceptance threshold[9]: only below this threshold offers are rejected. YNG responders accept or reject without knowing the offer. In either game, the responder's conflict payoff can be negative or positive and monitoring costs vary the efficiency advantage of trusting (Lewicki et al., 2006). Contrary to benchmark predictions, we expect YNG rather than UG to be more often selected by either player even when the responder's conflict payoff is positive and monitoring is costless. Nevertheless, we expected that which party has chosen the game affects game playing. Whereas responders may opt for YNG in the hope of being rewarded, may select YNG for the sake of "moral wiggle room", inexperienced participants can hardly anticipate how the game choice of the other party may signal its game playing[10].

We find that responder participants often opt for YNG but are not rewarded by offers which are at least as large as in UG (responder participants earn significantly more when opting for UG). Importantly, the outcomes of YNG and UG plays are at best weakly affected by who, proposer or responder, has chosen between games. This may change in future research, if it allows for learning in the light of feedback information about past results.

The paper is organized as follows: Section 2 introduces the games, how players elect one of them and describes the experimental design and protocols. Results are presented in Section 3. Section 4 concludes. The Appendix contains the Instructions (translated in English) and some additional tables.

[9] What is ruled out is, for instance, rejecting meager as well as overgenerous offers (see Güth and Kocher, 2014).
[10] Proposer participants, who choose YNG, do not "trust" but only the institution where "trust" is needed.

2. Endogenous game choice and experimental protocols

Both YNG and UG involve a proposer (*X*) and a responder (*Y*) and, in both games, *X* and *Y* can share the positive integer amount ("pie") p of money. The integer offer by proposer *X*, called *y*, is restricted to non-negative integers not exceeding p. We denote by c (≥ 0) the cost of monitoring, which is faced in UG but not in YNG, and by d the negative or positive conflict (or disagreement) payoff of responder *Y*. The agreement surplus is $p - d$ in YNG and $p - c - d$ in UG. When d is positive, it is assumed to exceed the *smallest* positive (integer) offer.

In YNG, responder *Y* chooses between acceptance $(\delta = 1)$ and rejection $(\delta = 0)$ of the unknown *y* with $0 \leq y \leq p$. Proposer *X* is the residual claimant whose conflict payoff is *0*. Given *X*'s offer, *X* earns $\delta(\, p + c - y)$ and *Y* receives $\delta y + (\, 1 - \delta)d$. If $d < 0$, the optimal choice of *Y* is $\delta^* = 1$ and an opportunistic proposer *X* will offer $y^* = 0$. For $\delta = 1$ in case of $d > 0$ proposer *X* would also offer $y = 0$ what renders δ* = 0 optimal for *Y*; so y*= 0 and δ* = 0 is the only (perfect) equilibrium.[11] Thus YNG with $p + c > d > 0$ is a social dilemma game whose agreement surplus $p + c - d$ is predicted to be lost.

In UG, responder *Y* chooses (experimentally) a monotonic acceptance strategy *Y via* a non-negative (integer) acceptance-threshold $\underline{y}$ with $0 \leq \underline{y} \leq p$ such that *Y* accepts, $\delta(y) = 1$, when $y \geq \underline{y}$, and *Y* rejects $\delta(y) = 0$ when $y < \underline{y}$. Proceeding by backward induction, *Y*'s optimal choice is $\underline{y}^* = 0$ when $d < 0$ but $\underline{y}^* = d$ or $d + 1$ when $d > 0$. Anticipating this, an opportunistic proposer *X* will choose $y^* = 0$ when $d < 0$ and $y^* = d$ or $d + 1$ when $d > 0$.

Behaviorally, we expected:

(I) for $d < 0$ nearly all responders opt for YNG; for larger c, a higher percentage of proposers opting for YNG

(II) even for $d > 0$, i.e. when the social dilemma aspect of YNG[12] suggests δ* = 0, some responders opting for YNG to signal trust; proposers to doubt that they will be trusted via $\delta = 1$ and therefore to select UG but - with larger c - an increasing percentage of proposers to opt for YNG.

Less related to our specific focus on "monitoring vs. trusting" we predict:

[11] There may exist other equilibrium outcomes with $\delta^* = 0$ and offers y in the range $0 \leq y < d$ for $d > 0$.

[12] Y's gift would be acceptance, $\delta = 1$, and X's gift a decent offer.

(III) less than usual altruistic punishment (i.e. no agreement) by responders Y, especially for $d < 0$ and when Y has selected UG.

By opting for UG in spite of $c > 0$, responders disclose their fear of exploitation. In the spirit of social preferences (see Cooper and Kagel, 2016) revealing suspicion and mistrust could induce (let-down averse) proposers to offer less and (fearful) responders to expect and accept less than usual in UG experiments without preceding endogenous game selection (see the review by Güth and Kocher, 2014).

Participants are randomly assigned to a role (proposer X, or responder Y) which they will keep during the experiment. Randomly matched pairs interact together across six rounds (or scenarios) with varying parameters p, c and d. Participants do not receive feedback on offers, their acceptances and acceptance thresholds between rounds and, most importantly, on the previous game choices of the other party. Therefore, all they know until the end of the experiment is their own game choices.

Four between-subject conditions (denoted 1, 2, 3 and 4) result from considering: (a) the sequence of monitoring cost c across rounds (increasing or decreasing) and (b) which role selects between UG and YNG (proposer X or responder Y). Table 1 provides the key aspects of the 2X2- factorial between-subjects design with the description of the values the parameters p, c, and d assume in each round.

Table 1. The 2X2-factorial design

Treatment	**Game choice by**	**Round 1**			**Round 2**			**Round 3**			**Round 4**			**Round 5**			**Round 6**		
		p	*c*	*d*	*p*	*c*	*d*	*p*	*c*	*d*	*p*	*c*	*d*	*p*	*c*	*d*	*p*	*c*	*d*
1	*X*	19	0	-2	19	0	2	21	4	-2	21	4	2	23	8	-2	23	8	2
2	*X*	23	8	-2	23	8	2	21	4	-2	21	4	2	19	0	-2	19	0	2
3	*Y*	19	0	-2	19	0	2	21	4	-2	21	4	2	23	8	-2	23	8	2
4	*Y*	23	8	-2	23	8	2	21	4	-2	21	4	2	19	0	-2	19	0	2

To balance the total stakes across the three cost levels, respectively c = 0, 4, 8, the pie size p is p = 19 for c = 0; p = 21 for c = 4; and p = 23 for c = 8. So, what can be shared in YNG varies from 19, 21 to 23 whereas in UG only p-c = 19, 17 or 15 is available. The total rewards of both games, p for YNG and p-c for UG, always add up to 38 but their difference depends on size c of monitoring cost. Total reward has to be distinguished from the agreement surplus, $p - d$ in YNG and $p - (c + d)$ in UG. In either game the agreement surplus, for d = -2 exceeds by 4 the agreement surplus for d = +2.

In total, 180 undergraduate students coming from the departments of Law, Political Science and Economics of LUISS Guido Carli University were recruited using Orsée (Greiner, 2015). No subject participated in more than one session. The software is based on z-Tree (Fischbacher, 2007). An experimenter read aloud the instructions. Before starting the experiment participants could privately ask for clarification. The average earning was 9.48 euro (plus 5 euro show up fee). The experiment lasted approximately one hour.

3. Results

Due to our focus on the institutional choice between trusting and monitoring, we first report results on proposers' and responders' choices between the two games. To disentangle the drivers behind these choices, we will present findings on offers, response behavior (acceptance rate in YNG and acceptance thresholds in UG) and on payoffs in both games.

Of course, the game choice by proposer or responder may be confounded with endogenous selection effects: responders opting for YNG, for example, may reveal or want to signal personal trust via opting for trusting, YNG, whereas responders opting for UG prefer to monitor their proposer's trustworthiness. Similarly, proposers may opt for UG to avoid their responder's suspicion whereas opting for YNG instead suggests tolerating the ambiguity of this institutional choice. As we do not the control treatment with an exogenous random choice between UG and YNG (see Footnote 10), we test for interaction effects of institutional choice and player role on game behavior. Theoretically, allowing one party to opt [13] for an institution (game) offers a screening potential. We stress this by varying conditions; for instance, when distinguishing which role, proposer or responder, decides between trusting or monitoring.

3.1 Endogenous game choice

Overall, trust is preferred to monitoring: participants choose YNG (UG) in 58.15% (41.85%) of all cases. Proposers self-servingly suggest trusting: when proposers make the game choice, YNG is selected in 73.11% of all cases. When responders choose, this occurs only in 43.84% of all cases ($Z=-9.741$, $p = 0.000$, two-tailed test, Wilcoxon rank-sum test - henceforth WRST); see Figure 1 for a graphical illustration. Although responder participants seem to anticipate that opting for UG will

[13] Initially allowing one of the interacting parties to opt for an institution is similar to principal agent theory, where usually principals are supposed to offer a menu of employment contracts allowing agents to choose one of them and thereby reveal their types. In case of commercial enterprises with many employees it seems natural that the principal, i.e. the employer, employs screening devices. But it is also possible that an agent applies screening: a very talented singular agent with an attractive outside option may confront the principal, for instance, with a menu of acceptable employment conditions, all inducing the agent not to choose the outside option. Similarly, one can easily imagine for our setup circumstances rendering it reasonable that the proposer, respectively the responder, confronts a menu of institutional options and thereby may reveal the own type. Of course, when appealing to commercial firms the case of proposers making the institutional choice as well as proposing how to share the pie seems more natural.

discourage low offers, many seem to prefer signalling their trust in proposer fairness via opting for YNG what should be encouraged additionally via larger costs of monitoring.

Result 1**.** *Proposer X (Responder Y) participants, when responsible for the game choice, opt for YNG significantly more (less) often than for UG.*

Figure 1: Game Choice: Shares of participants choosing YNG

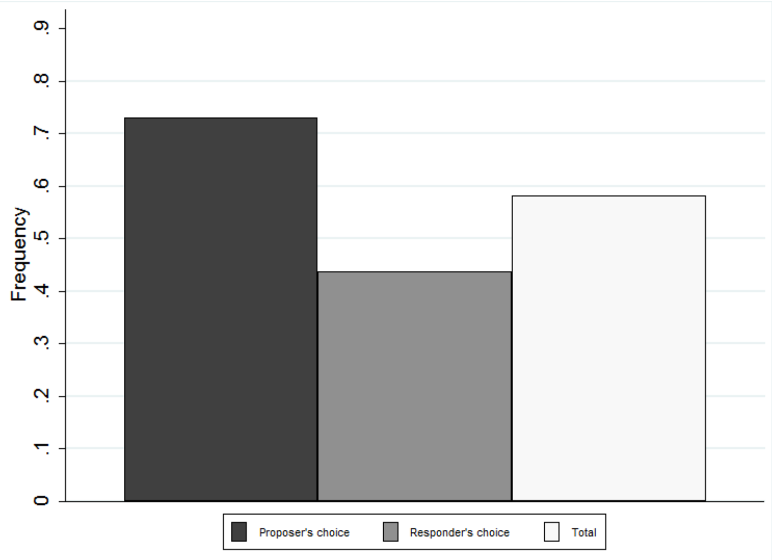


Table 2 presents YNG shares of participants who "predominantly" (in at least 4 of the 6 rounds) choose YNG: 75% of proposers predominantly choose YNG but only 32.6% of responders. On the other hand, 41.3% of responders predominantly choose UG but only 11.4% of proposers. A significant percentage of responders (26.1%) and proposers (13.6%) equally split their choice between the two games, i.e. are not predominantly opting for either one.

Table 2: (Relative) Frequency of participants who predominantly choose one game type

	Proposer (X)		Responder (Y)		Total	
	Freq.	Percent	Freq.	Percent	Freq.	Percent
Predominant UG	5	11.4 %	19	41.3%	24	26.7%
Mixed	6	13.6%	12	26.1%	18	20.0%
Predominant YNG	33	75.0%	15	32.6%	48	53.33%
Total	44	100%	46	100%	90	100%

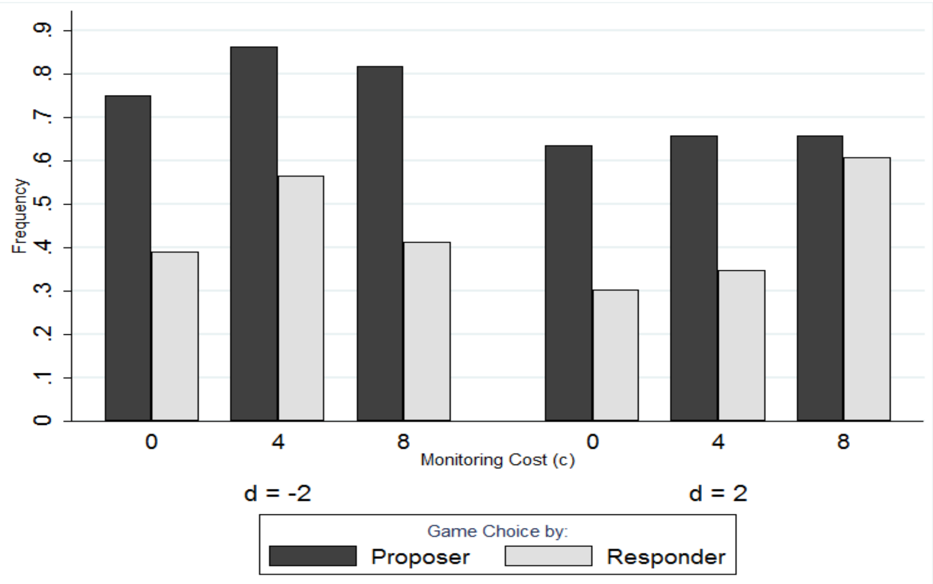
Note: Predominant = at least four of the six rounds.

The high general frequency of YNG choices (53.33%) in comparison to UG ones (26.7%) is far from obvious. Common sense suggests that proposers prefer not to be monitored and that responders prefer monitoring. In addition, suspicious proposer participants might fear rejection when self-servingly opting for YNG. Of course, one might also have opted for the game type with the larger payoff what can be controlled by $c = 0$.

3.1.1 Effects of monitoring costs (c) and their sequence

We expected participants to choose monitoring (UG) less frequently when $c > 0$ where zero monitoring costs do not justify trusting (YNG) by an efficiency motive. We indeed observe a significantly lower general frequency of YNG choices when $c = 0$ vs. $c > 0$ (51.66% vs. 61.38%, Z = -3.052, p = 0.002, WRST two-tailed test). This difference is driven by responders, who reduce their YNG percentage (from 48.37% for $c>0$ to 34.78% for $c=0$, Z = -3.030, p = 0.002, WRST two-tailed test) and thereby reveal their efficiency concerns. Proposers go in the same direction but only insignificantly (their YNG percentage reduction by $c = 0$ compared to $c > 0$ is 75% vs. 69.32%, WRST with Z = -1.387, p = 0.166, two-tailed test). One reading of these results could be that more responders than proposers want to signal efficiency concerns. Another explanation is that responders anyhow, i.e. in either game, are hoping that proposers are fairness minded and that inexperienced participants often underestimate or even neglect how monitoring enhances fairness concerns.

Figure 2: Detailed Game Choice – Relative Frequencies of selecting YNG



Conditions 1 and 3 differ from conditions 2 and 4 by presenting increasing vs. decreasing monitoring costs. There is no significant difference in game choices between increasing and decreasing costs *c* (for proposer: 70.63% vs. 75.36%, WRST with Z=1.222, p=0.222, two-tailed test; for responders 42.75% vs. 44.93%, WRST Z=0.514, p = 0.607, two-tailed test). Results by round and by increasing or decreasing costs *c* are reported in Table B1 in Appendix B. What apparently matters is not the sequence of cost levels but only the cost level *per se*: the correlation between YNG shares and monitoring cost level *c* is positive and highly significant (coeff.= 0.087, p=0.004, Pearson correlation test).

3.1.2 Effects of Y's conflict payoff being positive or negative

Comparing YNG-choices for $d = -2$ and $d = +2$ reveals (see Figure 2) that *d* does not significantly affect responders' game choices (54.35% vs. 57.97%, WRST with Z=0.857, p = 0.391, two-tailed test) but significantly enhances proposers' ones (81.06% vs. 65.15%, WRST with Z=4.118, p = 0.000, two-tailed test). Responders apparently neglect the social dilemma aspect of YNG for *d* = +2 for common(ly known) monetary opportunism due to $\delta{*}= 0$ and the possibility of receiving less than *d* = +2. Proposers, on the other hand, opt for trusting more often when *d* = -2.

Table 3: Determinants of game decision (YNG=1, UG=0).

	(1)	(2)	(3)	(4)
VARIABLES		interaction 1	interaction 2	round
Decision Maker (1 = Proposer)	0.786***	0.797***	0.928***	0.800***
	[0.147]	[0.141]	[0.199]	[0.147]
Positive Conflict Payoff (*d*) = 2	-0.275*	-0.094	-0.275**	-0.389***
	[0.142]	[0.166]	[0.127]	[0.151]
Monitoring cost dummies. Ref. category: c = 0				
Positive Monitoring Cost (*c*) = 4	0.248**	0.250**	0.316***	0.249**
	[0.105]	[0.099]	[0.084]	[0.104]
Positive Monitoring Cost (*c*) = 8	0.292**	0.292**	0.422***	0.306**
	[0.121]	[0.116]	[0.127]	[0.120]
Proposer as Decision Maker * *d*		-0.101*		
		[0.058]		
Proposer as Decision Maker * *c*			-0.036	
			[0.028]	
Round				0.108***
				[0.032]
Constant	-0.200*	-0.291**	-0.268*	-0.530***
	[0.131]	[0.137]	[0.149]	[0.136]
Observations	540	540	540	540

Probit (bootstrapped standard errors clustered at individual-level in brackets)

*** p<0.01, ** p<0.05, * p<0.1

Table 3 presents a set of probit regressions with bootstrapped standard errors. The table summarizes how experimental conditions affect the binary game choice between UG and YNG ("game decision"). It shows that the preference for YNG significantly increases across all specifications with (i) proposers making the game choice, (ii) higher monitoring costs (c), and (iii) over time. The results are consistent with previous non-parametric test results.

The non-parametric evidence suggested a significant effect of the conflict payoff *(d)* only when proposers choose between games: this effect is partially confirmed by the regression analysis in Table 3 (column 2). Furthermore, the significant and decreasing relationship between YNG-choices and conflict payoff (*d*) vanishes when controlling for the interaction between "Decision Maker" and "Conflict Payoff (*d*)", named "Proposer as Decision Maker * d". This interaction term is weakly significant.

Furthermore, column (3) controls for the interaction between "Decision Maker" and "Monitoring Cost (*c*)", named "Proposer as Decision Maker * c", which is not significant. Finally, column 4 controls for the round of interaction: the significant and positive effect suggests an increase in trusting of paired participants (in spite of not receiving any feedback information on previous rounds).

3.2 Proposer behavior

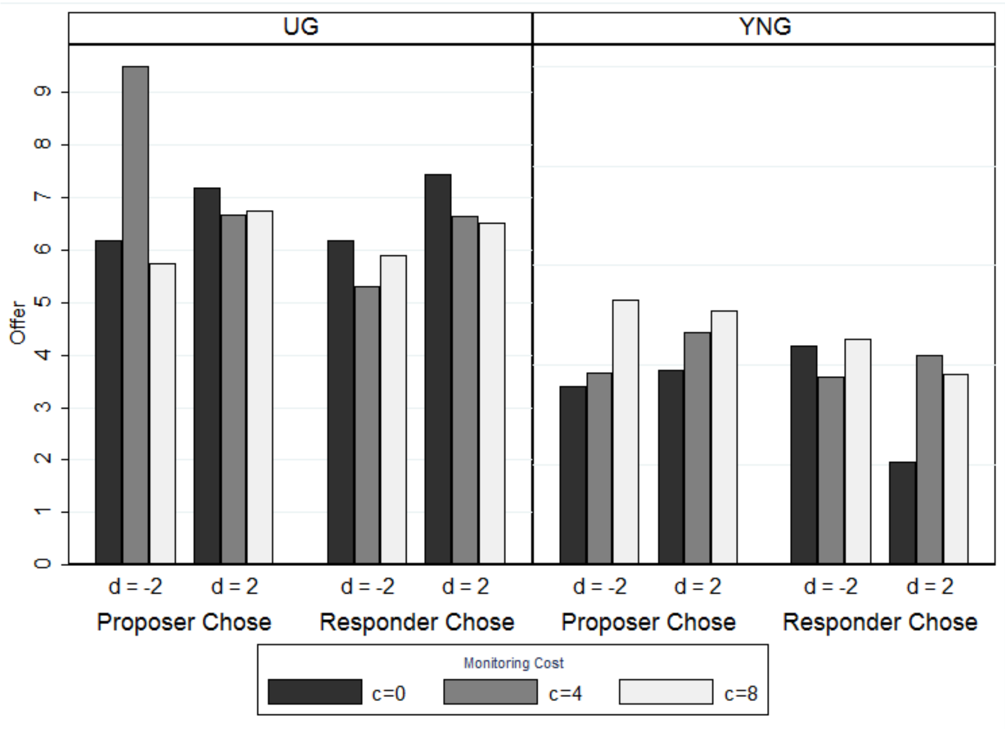
Do offers depend on the game, YNG or UG, and on who, *X* or *Y*, has chosen it? In the light of previous UG and YNG experiments, mostly based on $d = 0 = c$, we expected substantially and significantly larger UG than YNG offers, and both being larger for $d = +2$ than for $d = -2$ due to sharing the agreement surplus for $d = +2$. Our main results regarding offers confirm these expectations.

The average offer across all conditions is 5.17 corresponding to a 27.52% share of *p*. Proposers, in general, offer significantly more in UG than in YNG (6.54 vs. 4.18, WRST with Z=11.776, p = 0.000, two-tailed test). If we account for monitoring costs in UG, the difference in the percentage shares offered is even more striking: average offers amount to 38.34% of *p-c* in UG and 19.75% of *p* in YNG. Whether proposer or responder has opted for the game type does not significantly affect the offer level. Both when the proposer or the responder decided between games, we observe lower offers in YNG compared to UG. The difference is significant and substantial (proposers: 3.85 vs. 6.49, WRST with Z=7.759, p=0.000, two-tailed test; responders: 4.39 vs. 6.85, WRST with z=9.026, p=0.000, two-tailed test). These results are summarized in Figure 3[14].

Even in the special case $d = 0 = c$, on which earlier experiments in the literature are concentrated (see Güth and Kocher, 2014), UG offers are substantially and significantly larger than YNG ones (6.54 vs. 3.60, WRST with Z = 9.379, p = 0.000, two-tailed test).

[14] Table C1 (in Appendix C) shows the number of observations, mean, and standard deviation for each case represented in Figure 3.

Figure 3: Offers in UG vs. YNG



The average offer is significantly higher for d = + 2 than for d = - 2 (5.43 vs. 4.91, WRST with Z=-2.573, p = 0.010, two-tailed test). The correlation between offers in both games and monitoring costs is insignificant (coef.= 0.015, p=0.633, Pearson correlation test); restricted to UG-offers, we observe a negative and slightly significant correlation (coef.= -0.087, p=0.064, Pearson correlation test). The larger c, the more UG-interaction appears like a loss (due to not having opted for YNG) to what proposer participants apparently react by lower offers.

It is surprising that proposers do not reward responders who chose YNG by making them higher offers. When restricting the sample to responders opting for trust, UG-offers are still significantly higher (6.12 vs. 4.02 in YNG, WRST with Z = 6.524, p = 0.000, two-tailed test): responders, who opted for YNG, receive lower offers than those who have decided for UG, i.e. a responder favoring monitoring is, on average, better off than one favoring trust. Even worse, YNG-offers when the responder has made the game choice are significantly lower than when the proposer has made it (6.12 vs. 6.90, WRST with Z = 3.211, p = 0.001, two-tailed test). Rather than reciprocating responders' trust, proposers self-servingly opting for trust reveal their opportunism as intrinsically fairness minded proposers would not mind that offers are monitored.

Table 4 presents a set of GLS regressions with bootstrapped standard errors aimed at investigating the determinants of offer levels in both games. The main determinant of offer size is the chosen game: if YNG has been chosen (variable "YNG_chosen"), offers are significantly lower, no matter of the specification whereas proposer choosing the game (variable "Decision Maker (1 = Proposer)") does not significantly affect the offer.

Result 2. *Average UG-offers are significantly higher than YNG-ones, irrespective of whether proposers or responders have chosen the game. In both games, average offers are higher for d =+2 than for d =- 2 with this effect being not robust; average offers do not depend significantly on the players who chose the game and on the monitoring cost, c.*

Table 4: Offers

VARIABLES	(1)	(2) interaction 1	(3) interaction 2	(4) round
YNG Chosen	-3.154***	-3.125***	-3.435***	-3.073***
	[0.569]	[0.579]	[0. 808]	[0.544]
Decision Maker (1 = Proposer)	0.824	0.783	-0. 448	0.800
	[0.769]	[0.767]	[0.605]	[0.755]
Positive Conflict Payoff (*d*) = 2	0.266	0.840*	0.290	0.447
	[0.363]	[0.491]	[0.276]	[0.347]
Monitoring cost dummies. Ref. category: c = 0				
Positive Monitoring Cost (*c*) = 4	0.151	0.121	0.154	0.134
	[0.260]	[0.644]	[0.273]	[0.263]
Positive Monitoring Cost (*c*) = 8	0.422	0.450	0.436	0.389
	[0.297]	[0.330]	[0.333]	[0.316]
YNG Chosen * *d*		-0.251		
		[0.235]		
YNG * Proposer			0.290	
			[0.295]	
Round				-0.172**
				[0.088]
Constant	5.909***	5.602***	6.001***	6.400***
	[0.544]	[0.540]	[0.625]	[0.530]
Observations	540	540	540	540

Tobit Regressions with lower limit set to 0 (bootstrapped standard errors clustered at individual-level in brackets)
*** p<0.01, ** p<0.05, * p<0.1

The regression results generally confirm the other non-parametric test results. In particular, offers are higher for $d = +2$ than for $d = -2$, but with a weak effect which is not robust across all conditions. Controlling for interaction dummies of game choice and player role making this choice reveals no significant interaction effects on offers (column 3). Offers significantly decrease as interaction (round) proceeds (column 4) in spite of no information[15].

[15] Table B2 (in Appendix B) provides more details about offers, separately for treatments and rounds.

3.3 **Responder' behavior**

We now analyze response behavior, i.e. acceptance rate in YNG, and acceptance thresholds in UG. As anticipated, acceptance rates in YNG are generally high (85.35%); participants always accept in the first round, but acceptance rate decreases as the interaction proceeds (Spearman corr. test with coef.=-0.145, p=0.001) with 17% of rejections in rounds 2-6. The acceptance rate increases significantly from $d = +2$ to $d = -2$ (72.22% vs. 96.47%: WRST with Z=8.556, p =0.000, two-tailed test). Interestingly, the acceptance rate is not significantly affected by who has chosen YNG (84.97% vs. 85.95%: WRST with Z=0.336, p = 0.737, two-tailed test) or by the monitoring cost (85.52% vs. 84.94%: WRST with Z= -0.186, p = 0.852, two-tailed test). Table 5 reports a set of regressions investigating the determinants of the acceptance rate in YNG[16].

Table 5: Acceptance rate in the YNG

	(1)	(2)
VARIABLES		round
Decision Maker (1 = Proposer)	-0.063	-0.057
	[0.127]	[0.170]
Positive Conflict Payoff (*d*) = *0*	-1.261***	-1.206***
	[0.169]	[0.156]
Monitoring cost dummies. Ref. category: c = 0		
Positive Monitoring Cost (*c*) = 4	-0.157	-0.195
	[0.139]	[0.159]
Positive Monitoring Cost (*c*) = 8	0.184	0.166
	[0.169]	[0.200]
Round		-0.072*
		[0.042]
Constant	1.870***	2.130***
	[0.192]	[0.240]
Observations	628	628

Probit (bootstrapped standard errors, clustered at individual-level, in brackets)
*** p<0.01, ** p<0.05, * p<0.1

The regression analysis presented in Table 5 generally confirms the non-parametric results. The acceptance rate in YNG is significantly and negatively influenced by the conflict payoff, no matter of the specification. Whether proposer or responder chose between games does not play any role.

[16] Table B3 (in Appendix B) presents more detailed information on YNG acceptance rates.

Result 3. *Overall, YNG acceptance rate is significantly higher for d =-2 than for d =+2, without being significantly affected by who, X or Y, has chosen YNG or by the monitoring cost.*

Focusing now on UG-acceptance thresholds, we observe that the average threshold is 5.97 and becomes significantly smaller when *Y*'s outside option is negative (from 6.79 when *d* = +2 to 4.93 when *d* = -2: WRST with Z=-6.407, p = 0.000, two-tailed test). However, it also matters whether c is increasing or decreasing. The average threshold is 6.68 in treatments when *c* increases and only 5.22 when *c* decreases (WRST with Z=-4.869, p = 0.000, two-tailed test). In general, *c > 0* triggers a significantly lower average acceptance threshold (5.50) than for *c*= *0*, (6.71) (WRST with Z=-4.869, p = 0.000, two-tailed test).

Table 6 summarizes the result from a set of regressions on the determinants of the acceptance threshold in UG. They confirm the role of outside option, *d*, and monitoring costs, *c*, which is robust also when controlling for the round of interaction (see column 2). The acceptance threshold is lower when the responder has decided to play the Ultimatum Game.

Table 6: Acceptance threshold in UG

	(1)	(2)
VARIABLES	basics	time
Decision Maker (1 = Proposer)	-0.993*	-0.993*
	[0.577]	[0.559]
Positive Conflict Payoff (*d) = 2*	1.909***	1. 909***
	[0.213]	[0.270]
Monitoring cost dummies. Ref. category: c = 0		
Positive Monitoring Cost (*c*) = 4	-0.799***	-0. 799***
	[0.287]	[0.299]
Positive Monitoring Cost (*c*) = 8	-1.661***	-1.661***
	[0.273]	[0.291]
Round		0.001
		[0.083]
Constant	5.968***	5.966***
	[0.385]	[0.359]
Observations	452	452

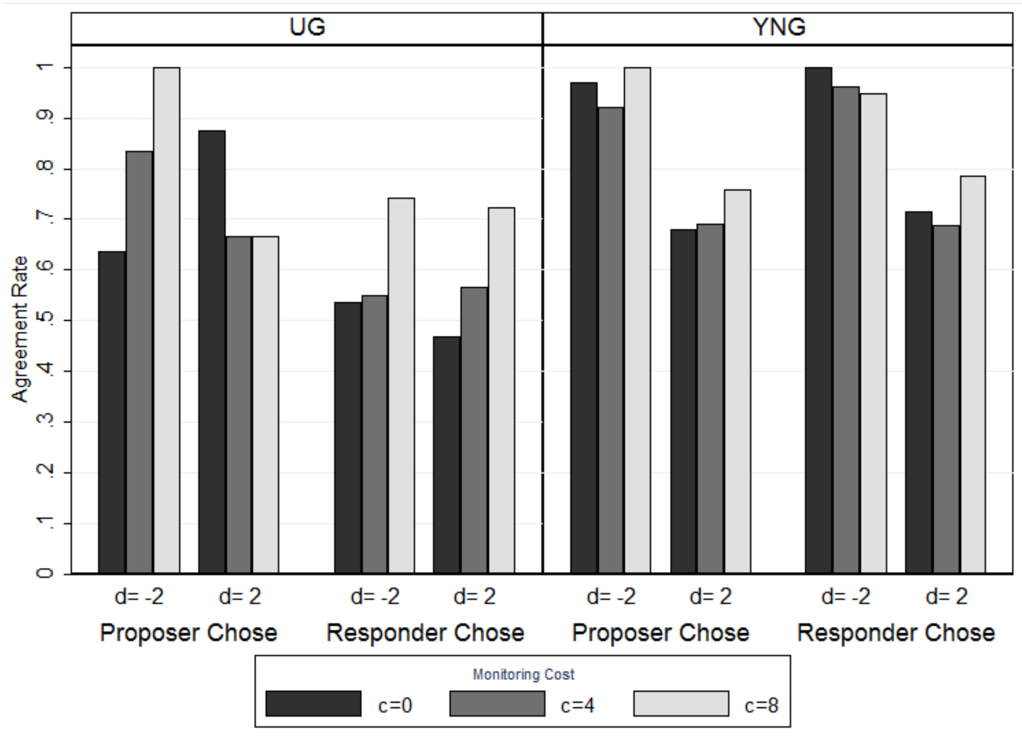
GLS (bootstrapped standard errors clustered at individual-level in brackets)

*** p<0.01, ** p<0.05, * p<0.1

Result 4. *The average acceptance threshold in UG is significantly smaller for d = -2 than d = +2 and decreases with increasing cost of monitoring c. Acceptance in UG is more likely for lower monitoring costs and when the responder is the one who has chosen the game.*

Figure 4[17] reports agreement rates by distinguishing between game types, who has selected them, the conflict payoff, and the monitoring cost. Agreements are significantly more frequent in YNG than UG (85.35% vs. 64.15%: WRST with Z=-8.907, p =0.000, two-tailed test), for strictly positive monitoring costs (78.61% vs. 72.22%: WRST with Z=2.333, p=0.019, two-tailed test), and when proposers have chosen the game (82.58% vs. 70.65%, WRST with Z=4.616 and p=0.000, two-tailed test). The responder's conflict payoff plays an important role as well: negative conflict payoff, *d =-2*, leads to significantly more agreements (85.18% vs. 67.78%, WRST with Z= 6.741, p=0.000, two-tailed test).

Figure 4: Agreement rates



Result 5**.** *Agreement is significantly more frequent in YNG than in UG and more likely when proposers have chosen the game, when the monitoring is costly, and when Y's conflict payoff is negative.*

[17] Table C2 (in Appendix C) shows the number of observations, mean, and standard deviation for each case represented in Figure 4.

Table B3 in Appendix B presents agreement rates for all the rounds and treatments in YNG, and Tables B4 and B5 present acceptance thresholds and agreement rates for all rounds and treatments in UG.

3.4 Payoffs

Since in terms of acceptance rates YNG outperforms UG but in YNG offers are smaller, responders might, on average be better off in YNG than in UG. Table 7 displays the average payoffs in both games for proposers and responders, separately for who has chosen the game type.

Table 7: Average Payoff by Game Type and Role Selecting Between Games

		Proposer (X)		Responder (Y)		Sum (X + Y)	
		UG	YNG	UG	YNG	UG	YNG
Decision Maker	X	7.52	13.99	5.69	4.14	13.21	18.13
	Y	5.47	14.80	4.55	3.66	10.02	18.46

Responders choosing YNG on average make the "wrong" game choice and lose 0.89 euro (4.55-3.66) whereas the corresponding gain is of 6.47 euro (13.99 - 7.52) for proposers. Both differences are significant (WRTS with Z=-14.138 and p=0.000 for proposers, and Z=5.232 and p=0.000, two-tailed test). Assuming "true" (so-called "rational") expectations, proposers opting for UG and responders opting for YNG choose the worse game. The loss for proposers in choosing monitoring decreases when monitoring costs are 0 but remains highly significant (see Table B6 in Appendix B).

Result 6. *According to average payoffs, responders should avoid YNG (trust) whereas proposer participants should prefer it.*

4. Conclusion

Deciding between trusting and monitoring is crucial in many private relationships (e.g. husband and wife or child and parent) as well as in economic interactions like market and intra-firm exchanges. In case of market deliveries, customers may want to check product quality or trust the seller and whether and how to monitor is a crucial aspect of corporate governance. Addressing this issue is thus important.

We were interested in what screening via an endogenous institutional choice of trusting or monitoring implies for game playing and its outcome, and not in comparing treatments differing in only one controlled aspect. Specifically, our experiment asks participants to decide between "Trusting" vs. "Monitoring" via selecting the Yes-No-Game or the Ultimatum Game. This possibly confounds the effects of the chosen game with what the choosing party, proposer or responder, may reveal by its choice. [18] In our view, anticipating such confounding effects by inexperienced participants is possible but rather far-fetched in a complex experimental setup as ours which could be confirmed by our regression results. But, of course, one might also have run the control treatment which rules out screening.

Surprisingly only 23.9% of responders predominantly prefer to monitor, i.e. opt for YNG which, on average, let them earn less. This can, of course, persist due to no feedback on previous outcomes like the other's previous game choices what excludes learning. The predominant preference for YNG, i.e. for trusting, thus so far applies only to newly formed dyads who interact just once, for instance, as traders on spot markets. Whether responders would learn to prefer UG over YNG, i.e. substitute trusting by monitoring, would be interesting to explore in future research allowing learning based on feedback information about past outcomes.[19]

Another surprising result is that. proposers self-servingly select trusting (YNG) even when the cost of monitoring is zero. And they are perfectly right doing so: no single YNG-offer has been rejected in the first round, only with more familiarity, a significantly positive tendency to reject in YNG has evolved. Average offers, acceptance rates (in YNG) and acceptance thresholds (in UG) are qualitatively in line with usual results, albeit tending to the meager side. Monitored UG-offers, on average, significantly exceed YNG-ones without preventing a significantly positive and surprisingly large UG rejection rate. Game offers and acceptance rates are hardly influenced by who, *X* or *Y*, has opted for the game (YNG or UG) what supports our intuition that in a rather complex setup anticipating screening effect or signaling the (motivational) type by the party, who has opted for the game, is unlikely for inexperienced participants

[18] When participants are aware of the institutional alternative, behavior might depend not only on the given institution but also on its (foregone) alternative chosen what is captured by random selection of one institution, UG or YNG.

[19] The predominance of trusting might erode when providing feedback information which, in turn, might affect the fairness of YNG-offers especially for positive cost of monitoring ($c > 0$). So far our conclusions inform on initial and intuitive tendencies on how to resolve the conflict between trusting, when doubting trustworthiness, and costly monitoring.

References

Audretsch, D. B., and Thurik, A. R. (2000). 'Capitalism and Democracy in the 21st Century: From the Managed to the Entrepreneurial Economy', Journal of Evolutionary Economics 10(1), 17–34.

Audretsch, D. B., and Thurik, A. R. (2001). 'What is New about the New Economy: Sources of Growth in the Managed and Entrepreneurial Economies', Industrial and Corporate Change 10, 267–315.

Avenhaus, R. (2004). "Application of inspection games", Mathematical Modelling and Analysis, 9 (3), 179-192.

Benabou, R., and Tirole, J. (2003). Intrinsic and extrinsic motivation. Review of Economic Studies, 70, 489-520.

Berg, J., Dickhaut, J., and McCabe, R. (1995). Trust, reciprocity and social history. Games and Economic Behavior, 10 (1), 122-142.

Bijlsma-Frankema, K., de Jong, B., and van de Bunt, G. (2008). Heed, a missing link between trust, monitoring and performance in knowledge-intensive teams. The International Journal of Human Resource Management, 19 (1), 19-40.

Bowles, S., and Polania-Reyes, S. (2012). Economic incentives and social preferences: substitutes or complements? Journal of Economic Literature, 50 (2), 368-425.

Boxall, P., 1996. 'The Strategic HRM Debate and the Resource-Based View of the Firm', HRM Journal 6(3), 59–75.

Calvo, G.A., and Wellisz, S. (1978). Supervision, Loss of Control, and the Optimum Size of the Firm. Journal of Political Economy, 86, 943-52.

Cooper, D.J., and Kagel, J.H. (2016). Other-regarding preferences: A selective survey of experimental results, in The Handbook of Experimental Economics, Vol. 2, J.H. Kagel and A.E. Roth (eds), Princeton University Press.

Dana, J, Weber, R.A. and Kuang, J.X., (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness, Economic Theory, 33 (1), 67-80.

De Jong, B.A., and Elfring, T. (2010). How does trust affect the performance of ongoing teams? The mediating role of reflexivity, monitoring, and effort. Academy of Management Journal, 53(3), 535-549.

De Kok, J., and Uhlaner, L. M. (2001). Organization context and human resource management in the small firm. Small business economics, 17(4), 273-291.

Dickinson, D., and Villeval, M. C. (2008). Does monitoring decrease work effort?: The complementarity between agency and crowding-out theories. Games and Economic Behavior, 63(1), 56-76.

Dirks, K.T., and Ferrin, D.L., (2001). The role of trust in organizational settings. Organization Science, 12(4), 450-467.

Falk, A., Gächter, S., and Kovacs, J. (1999). Intrinsic motivation and extrinsic incentives in a repeated game with incomplete contracts. Journal of Economic Psychology, 20, 251-284.

Falk, A. and Kosfeld, M., (2006). The hidden cost of control. American Economic Review, 96, 1611-1630.

Ferrin, D.L., Bligh, M.C., and Kohles, J.C. (2007). Can I trust you to trust me? A theory of trust, monitoring, and cooperation in interpersonal and intergroup relationships. Group and Organization Management, 32(4), 465-499.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments, Experimental Economics, 10 (2), 171-178.

Frey, B. (1993). Does monitoring increase work effort? The rivalry between trust and loyalty. Economic Inquiry, 31, 663-670.

Fullerton, D. and Metcalf, G.E. (2002). Tax incidence. Handbook of public economics. Elsevier.

Gächter, S., and Falk, A., (2002). Work motivation, institutions, and performance, in R. Zwick and A. Rapoport, Experimental Business Research. Boston: Kluwer Academic Publishers, 351-72.

Greiner, B. (2015). Subject pool recruitment procedure: Organizing experiments with ORSEE. Journal of the Economic Science Association, 1 (1), 114-125.

Güth, W., and Kocher, M.G. (2014). More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. Journal of Economic Behavior and Organization, 108, 396-409.

Hendrikse, G., Hippmann, P., & Windsperger, J. (2015). Trust, transaction costs and contractual incompleteness in franchising. Small Business Economics, 44(4), 867-888.

Kim, P.H., Dirks, K.T., Cooper, C.D., and Ferrin, D.L., (2006). When more blame is better than less: The implications of internal vs. external attributions for the repair of trust after a competence- vs. integrity-based trust violation. Organizational Behavior and Human Decision Processes, 99(1), 49-65.

Langfred, C. W. (2004). Too much of a good thing? Negative effects of high trust and individual autonomy in self-managing teams. Academy of management journal, 47(3), 385-399.

Lewicki, R.J., Tomlinson, E.C., and Gillespie, N., (2006). Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions. Journal of Management, 32(6), 991-1022.

Malhotra, D., (2004). Trust and reciprocity decisions: The differing perspectives of trustors and trusted parties. Organizational Behavior and Human Decision Processes, 94(2), 61-73.

Nagin, D., Rebitzer, J., Sanders, S., and Taylor, L. (2002). Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. American Economic Review, 92, 850-873.

Prendergast, C. (1999). The Provision of Incentives in Firms. Journal of Economic Literature, 37(1), 7-63.

Schweitzer, M., Ho, T., and Zhang, X. (2016) How monitoring influences trust: A tale of two faces. Management Science, 64 (1), 253-270.

Appendix A: Instructions (Treatment 1)

Welcome and thank you for participating in this experiment. You are participating in a study on economic decision making. During the experiment you can, depending on your decisions and on other participants' decisions, earn considerably more money in addition to the 5 euros for showing-up on time. Your answers and your choices will be totally anonymous. The experimenters will not be able to associate your choices and your answers to your name.

During the experiment, you cannot communicate with other participants (otherwise you would be excluded from the experiment) and you should be very careful in reading the instructions that will appear on your screen and will be read out by one of the experimenters. If you have any questions, please raise your hand and an experimenter will come to individually answer your questions.

At the end of the experiment, you will be asked to fill a short questionnaire; afterwards, you will be privately paid.

Participants will be matched into pairs, with one member assigned to role X and the other to role Y. Roles will be assigned randomly. The role assigned and the pair composition will remain constant for the whole experiment. Therefore, you will always interact with the same person, whose identity will not be revealed to you. Similarly, your identity will not be revealed to your partner.

In each pair, X and Y have to decide how to split an amount of money in six scenarios that will be described in detail during the experiment. You will not receive any feedback on previous outcomes between scenarios. Only one of the six scenarios will be randomly selected and actually paid, with each of them being equally probable.

In each scenario, the participants in role X have to choose between two mechanisms: mechanism M and mechanism T.

In mechanism M, X and Y can share a given amount of euros. Role X decides how many euros to keep, and how many euros to offer to Y. Y has to state, before knowing X's offer an acceptance threshold below which X's offer will be refused. If Y refuses the offer, X will obtain zero euros and Y will earn or lose two euros depending on the specific scenario. If Y accepts the offer, i.e. when the offered amount to Y is not lower than Y's acceptance threshold, Y receives the offer and X receives the amount kept.

In mechanism T, X and Y can share a given amount of euros. Role X decides how many euros to keep for, and how many euros to offer to Y. Y has to decide whether to accept any offer (without knowing what the offer is). If Y refuses the offer, X will obtain zero euros and Y will earn or lose two euros depending on the specific scenario. If Y accept. Y receives the offer and X receives the amount kept.

The 6 scenarios differ in Y's gain or loss of two euros if Y rejects, and the possibly different euro amounts to be split in the two mechanisms. Before X decides on the mechanism, X and Y will, both, be informed on the screen about these parameters.

Scenario 1

X has to decide if the pair will play mechanism M or mechanism T.

Mechanism M: X and Y can share 19 euros and role X decides how many euros to keep for how many euros to offer to Y. Role Y has to state the acceptance threshold below which he/she will refuse X's offer. If X's offer is refused, X will earn zero euros and Y will lose 2 euros.

Mechanism T: X and Y can share 19 euros and role X decides how many euros to keep and how many euros to offer to Y. Role Y decides whether to accept or refuse any offer. If X's offer is refused, X will earn zero euros and Y will lose 2 euros.

Scenario 2

X has to decide if the pair will play mechanism M or mechanism T.

Mechanism M: X and Y can share 19 euros and role X decides how many euros to keep for how many euros to offer to Y. Role Y has to state the acceptance threshold below which he/she will refuse X's offer. If X's offer is refused, X will earn zero euros and Y will earn 2 euros.

Mechanism T: X and Y can share 19 euros and role X decides how many euros to keep and how many euros to offer to Y. Role Y decides whether to accept or refuse any offer. If X's offer is refused, X will earn zero euros and Y will earn 2 euros.

Scenario 3

X has to decide if the pair will play mechanism M or mechanism T.

Mechanism M: X and Y can share 17 euros and role X decides how many euros to keep for how many euros to offer to Y. Role Y has to state the acceptance threshold below which he/she will refuse X's offer. If X's offer is refused, X will earn zero euros and Y will lose 2 euros.

Mechanism T: X and Y can share 21 euros and role X decides how many euros to keep and how many euros to offer to Y. Role Y decides whether to accept or refuse any offer. If X's offer is refused, X will earn zero euros and Y will lose 2 euros.

Scenario 4

X has to decide if the pair will play mechanism M or mechanism T.

Mechanism M: X and Y can share 17 euros and role X decides how many euros to keep for how many euros to offer to Y. Role Y has to state the acceptance threshold below which he/she will refuse X's offer. If X's offer is refused, X will earn zero euros and Y will earn 2 euros.

Mechanism T: X and Y can share 21 euros and role X decides how many euros to keep and how many euros to offer to Y. Role Y decides whether to accept or refuse any offer. If X's offer is refused, X will earn zero euros and Y will earn 2 euros.

Scenario 5

X has to decide if the pair will play mechanism M or mechanism T.

Mechanism M: X and Y can share 15 euros and role X decides how many euros to keep for how many euros to offer to Y. Role Y has to state the acceptance threshold below which he/she will refuse X's offer. If X's offer is refused, X will earn zero euros and Y will lose 2 euros.

Mechanism T: X and Y can share 23 euros and role X decides how many euros to keep and how many euros to offer to Y. Role Y decides whether to accept or refuse any offer. If X's offer is refused, X will earn zero euros and Y will lose 2 euros.

Scenario 6

X has to decide if the pair will play mechanism M or mechanism T.

Mechanism M: X and Y can share 15 euros and role X decides how many euros to keep for how many euros to offer to Y. Role Y has to state the acceptance threshold below which he/she will refuse X's offer. If X's offer is refused, X will earn zero euros and Y will earn 2 euros.

Mechanism T: X and Y can share 23 euros and role X decides how many euros to keep and how many euros to offer to Y. Role Y decides whether to accept or refuse any offer. If X's offer is refused, X will earn zero euros and Y will earn 2 euros.

[Treatment 2 – Scenario order: 5,6,3,4,1,2

Treatments 3 & 4 – As 1 and 2 but Responder chooses game]

Appendix B: Additional Descriptives

Table B1. Game decision by round

	Choice	Round 1	2	3	4	5	6	Total
X, c↗	UG	38.1%	47.6%	19.0%	38.1%	4.8%	28.6%	29.4%
	YNG	61.9%	52.4%	81.0%	61.9%	95.2%	71.4%	70.6%
X, c↘	UG	30.4%	39.1%	8.7%	30.4%	13.0%	26.1%	24.6%
	YNG	69.6%	60.9%	91.3%	69.6%	87.0%	73.9%	75.4%
Y, c↗	UG	69.6%	69.6%	47.8%	69.6%	52.2%	34.8%	57.2%
	YNG	30.4%	30.4%	52.2%	30.4%	47.8%	65.2%	42.8%
Y, c↘	UG	65.2%	43.5%	39.1%	60.9%	52.2%	69.6%	55.1%
	YNG	34.8%	56.5%	60.9%	39.1%	47.8%	30.4%	44.9%

Table B2. Offers by Game Decision

Round		X, c↗ UG	X, c↗ YNG	X, c↗ Total	X, c↘ UG	X, c↘ YNG	X, c↘ Total	Y, c↗ UG	Y, c↗ YNG	Y, c↗ Total	Y, c↘ UG	Y, c↘ YNG	Y, c↘ Total
1	Mean	5,75	3,85	4,57	5,71	6,63	6,35	6,13	2,57	5,04	6,20	8,13	6,87
	Std. Dev.	4,33	4,54	4,46	2,43	4,63	4,05	2,85	3,55	3,43	1,97	5,44	3,57
	Freq.	8	13	21	7	16	23	16	7	23	15	8	23
2	Mean	7,50	4,45	5,90	6,22	5,50	5,78	7,44	1,43	5,61	6,70	4,46	5,43
	Std. Dev.	1,78	6,22	4,82	0,97	4,03	3,18	2,56	1,81	3,65	1,34	5,21	4,10
	Freq.	10	11	21	9	14	23	16	7	23	10	13	23
3	Mean	10,25	3,29	4,62	8,00	4,29	4,61	5,73	1,33	3,43	4,78	5,86	5,43
	Std. Dev.	4,72	3,53	4,60	1,41	3,15	3,20	3,10	1,92	3,36	5,04	5,83	5,44
	Freq.	4	17	21	2	21	23	11	12	23	9	14	23
4	Mean	6,63	4,23	5,14	6,71	5,00	5,52	6,19	3,86	5,48	7,14	4,44	6,09
	Std. Dev.	2,67	5,37	4,61	2,50	3,58	3,33	2,32	3,89	3,00	2,66	6,91	4,83
	Freq.	8	13	21	7	16	23	16	7	23	14	9	23
5	Mean	6,00	4,25	4,33	7,33	3,40	3,91	5,50	1,91	3,78	6,25	5,55	5,91
	Std. Dev.	0,00	5,08	4,96	2,08	2,60	2,84	3,87	4,11	4,31	3,62	5,54	4,54
	Freq.	1	20	21	3	20	23	12	11	23	12	11	23
6	Mean	7,50	4,73	5,52	6,67	3,53	4,35	6,25	3,27	4,30	7,44	2,71	6,00
	Std. Dev.	2,88	5,05	4,64	2,80	3,24	3,38	2,43	4,96	4,44	3,01	2,43	3,57
	Freq.	6	15	21	6	17	23	8	15	23	16	7	23
Total	Mean	7,19	4,11	5,02	6,50	4,63	5,09	6,27	2,39	4,61	6,54	5,24	5,96
	Std. Dev.	3,24	4,82	4,62	2,08	3,62	3,40	2,85	3,69	3,76	3,04	5,48	4,34
	Freq.	37	89	126	34	104	138	79	59	138	76	62	138

Table B3. Acceptance rate in YNG

Round		Treatment				Total
		X, c↗	**X, c↘**	**Y, c↗**	**Y, c↘**	
1	Mean	1.00	1.00	1.00	1.00	1.00
	Std. Dev.	0.00	0.00	0.00	0.00	0.00
	Freq.	26	32	14	16	88
2	Mean	0.82	0.79	0.86	0.69	0.78
	Std. Dev.	0.39	0.42	0.36	0.47	0.42
	Freq.	22	28	14	26	90
3	Mean	0.88	0.95	1.00	0.93	0.94
	Std. Dev.	0.33	0.22	0.00	0.26	0.24
	Freq.	34	42	24	28	128
4	Mean	0.62	0.75	0.86	0.56	0.69
	Std. Dev.	0.50	0.44	0.36	0.51	0.47
	Freq.	26	32	14	18	90
5	Mean	1.00	0.95	0.91	1.00	0.97
	Std. Dev.	0.00	0.22	0.29	0.00	0.18
	Freq.	40	40	22	22	124
6	Mean	0.73	0.59	0.87	0.57	0.70
	Std. Dev.	0.45	0.50	0.35	0.51	0.46
	Freq.	30	34	30	14	108
Total	Mean	0.85	0.85	0.92	0.81	0.85
	Std. Dev.	0.35	0.36	0.28	0.40	0.35
	Freq.	178	208	118	124	628

Table B4. Acceptance threshold in UG

Round		Treatment X, c↗	X, c↘	Y, c↗	Y, c↘	Total
1	Mean	4,75	1,86	7,13	3,80	4,83
	Std. Dev.	3,92	1,51	1,72	2,78	3,15
	Freq.	16	14	32	30	92
2	Mean	7,00	5,11	8,50	5,30	6,78
	Std. Dev.	3,84	2,81	2,87	2,87	3,38
	Freq.	20	18	32	20	90
3	Mean	6,50	3,50	5,73	4,56	5,27
	Std. Dev.	5,04	0,58	2,31	3,22	3,15
	Freq.	8	4	22	18	52
4	Mean	7,63	4,86	6,44	6,93	6,56
	Std. Dev.	3,50	3,57	2,78	2,42	3,02
	Freq.	16	14	32	28	90
5	Mean	5,00	3,67	4,67	5,17	4,79
	Std. Dev.	0,00	4,13	2,44	5,02	3,85
	Freq.	2	6	24	24	56
6	Mean	7,33	5,17	7,38	7,63	7,11
	Std. Dev.	3,89	3,16	1,36	2,73	2,89
	Freq.	12	12	16	32	72
Total	Mean	6,59	4,18	6,72	5,68	5,97
	Std. Dev.	3,93	3,09	2,65	3,48	3,34
	Freq.	74	68	158	152	452

Table B5: Agreements rate across rounds in UG

	X, c↗		**X, c↘**		**Y, c↗**		**Y, c↘**		**Total**	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
Rejection	22	29,7%	12	17,6%	74	46,8%	54	35,5%	162	35,8%
Acceptance	52	70,3%	56	82,4%	84	53,2%	98	64,5%	290	64,2%
Total	74	100%	68	100%	158	100%	152	100%	452	100%

Table B6: Average Payoff by Monitoring Cost

		Proposer (X)		Responder (Y)		Sum (X + Y)	
Average Payoff (c=0)							
		UG	YNG	UG	YNG	UG	YNG
	X	9.18	12.44	4.55	3.7	13.73	16.14
Game Chosen by	Y	5.12	13.56	4.52	3.31	9.64	16.87
Average Payoff (c>0)							
		Proposer (X)		Responder (Y)		Sum (X + Y)	
		UG	YNG	UG	YNG	UG	YNG
	X	6.5	14.71	5.84	4.34	12.34	19.05
Game Chosen by	Y	5.69	15.24	4.57	3.79	10.26	19.03

Table B7: Average Payoff by Conflict Payoff

		Proposer (X)		Responder (Y)		Sum (X + Y)	
Average Payoff (d=-2)							
		UG	YNG	UG	YNG	UG	YNG
	X	7.76	16.21	5.36	4.01	13.12	20.22
Game Chosen by	Y	5.87	16.16	3.65	4.11	9.52	20.27
Average Payoff (d=2)							
		Proposer (X)		Responder (Y)		Sum (X + Y)	
		UG	YNG	UG	YNG	UG	YNG
	X	7.39	11.24	5.86	4.3	13.25	15.54
Game Chosen by	Y	5.1	13.33	5.39	3.17	10.49	16.50

Appendix C – Figure complements

Table C1 - Mean offers and Std. Dev. for each case (compliments Figure 3)

DM	*d*	*c*	UG			YNG		
			N	Mean	Std. Dev.	N	Mean	Std. Dev.
Proposer	-2	0	22	6.182	3.724	66	3.576	3.411
	2	0	32	7.188	2.132	56	3.893	4.515
	-2	4	12	9.500	3.705	76	3.842	3.295
	2	4	30	6.667	2.454	58	4.655	4.363
	-2	8	16	5.750	2.176	72	5.306	4.927
	2	8	30	6.733	1.946	58	5.103	4.483
Responder	-2	0	56	6.179	3.111	36	4.389	4.901
	2	0	64	7.438	2.725	28	2.071	2.124
	-2	4	40	5.300	3.950	52	3.769	4.909
	2	4	60	6.633	2.463	32	4.188	5.533
	-2	8	54	5.889	2.899	38	4.526	5.476
	2	8	36	6.500	1.828	56	3.821	4.973

Table C2 - Mean acceptance rates and Std. Dev. for each case (compliments Figure 4)

DM	*d*	*c*	UG			YNG		
			N	Mean	Std. Dev.	N	Mean	Std. Dev.
Proposer	-2	0	22	0.636	0.492	66	0.970	0.173
	2	0	32	0.875	0.336	56	0.679	0.471
	-2	4	12	0.833	0.389	76	0.921	0.271
	2	4	30	0.667	0.479	58	0.690	0.467
	-2	8	16	1.000	0.000	72	1.000	0.000
	2	8	30	0.667	0.479	58	0.759	0.432
Responder	-2	0	56	0.536	0.503	36	1.000	0.000
	2	0	64	0.469	0.503	28	0.714	0.460
	-2	4	40	0.550	0.504	52	0.962	0.194
	2	4	60	0.567	0.500	32	0.688	0.471
	-2	8	54	0.741	0.442	38	0.947	0.226
	2	8	36	0.722	0.454	56	0.786	0.414